DOCUMENT RESUME

ED 054 365 VT 013 759

TITLE	Training Evaluation: A Guide to Its Planning, Development, and Use in Agency Training Courses. Training Systems and Technology Series 4.
INSTITUTION	Civil Service Commission, Washington, D.C. Bureau of Training.
REPORT NO	Pamph-T-13
PUB DATE	May 71
NOTE	21p.
AVAILABLE FROM	Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402 (CS1.48:T13, $.30)
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Administrator Guides, Conceptual Schemes, Data Collection, *Evaluation Methods, *Inplant Programs, Planning, Program Descriptions, *Program Evaluation, *Systems Approach, Training Objectives, Vocational Followup

ABSTRACT

This paper describes an approach which can be used by the agency training specialist to assess the effectiveness of many internally developed and conducted training courses. It offers a conceptual framework for evaluation, the essential steps in developing training objectives, an in-course training evaluation process, and a description of the steps involved in developing and using data gathering instruments. The paper concludes with a discussion of the factors to be considered when implementing this type of evaluation program. (Author/JS)

ED054365

Training Evaluation:
A Guide to its Planning, Development, and Use in Agency Training Courses

Training Systems and Technology Series: No. IV

U.S. CIVIL SERVICE COMMISSION
BUREAU OF TRAINING

VT013759

TRAINING SYSTEMS AND TECHNOLOGY SERIES PUBLICATIONS

No. I Instructional Systems and Technology: An Introduction to the Field and Its Use in Federal Training (Pamphlet T–11). Price—30 cents.

No. II Application of a Systems Approach to Training: A Case Study (Pamphlet T–2). Price—30 cents.

No. III Programmed Instruction: A Brief of its Development and Current Status (Pamphlet T–12). Price—55 cents.

No. IV Training Evaluation: A Guide to Its Planning, Development, and Use in Agency Training Courses (Pamphlet T–13). Price 30 cents.

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402
Price 30 cents

ED054365

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

PAMPHLET T-13

Training Evaluation: A Guide to Its Planning, Development, and Use in Agency Training Courses

MAY 1971

U.S. Civil Service Commission
Bureau of Training
Training Systems and
Technology Division

TABLE OF CONTENTS

	Page
Introduction	1
Conceptual Framework for Training Course Evaluation	1
Developing Training Objectives and Standards	2
An In-Course Training Evaluation Process	3
Direct Data about Trainee Achievement	5
Indirect Data about Trainee Achievement	5
Developing and Using Direct Measurement Instruments	6
Developing and Using Indirect Measurement Instruments	10
Factors That Influence the Decision To Evaluate	15
Footnotes	16
Selected Bibliography	16

INTRODUCTION

Training Evaluation: A Guide to Its Planning, Development, and Use in Agency Training Courses is the fourth in a series being published by the Bureau of Training to examine various aspects of training technology and the application of an instructional systems approach to agency training programs. An instructional systems approach is simply a way of dealing with organizational training requirements by accurately identifying the organization's problems and needs, setting specific objectives, applying logic and analysis techniques to the problems, developing methods for their solution, and comparing measurement results with the established objectives.[1] The growing need for more effective and efficient training has prompted increasing numbers of training specialists and management officials to consider this analytical approach in the design, development, and assessment of their training courses.

When applied to the assessment of training courses, a systems approach can provide continuing feedback of trainee performance data for use by the training staff in judging the effectiveness of their training. In this context, the process by which these data are collected, analyzed, and interpreted is training evaluation.

This paper describes one approach which can be used by the agency training specialist to assess the effectiveness of many of their internally developed and conducted training courses. It offers a conceptual framework for evaluation; the essential steps in developing training objectives; an in-course training evaluation process; and a description of the steps involved in developing and using data gathering instruments. The paper concludes with a discussion of the factors to be considered when implementing the kind of evaluation program described in this paper. However, the complexity of the total evaluation process precludes complete coverage here. Additional papers will be issued periodically supplementing and expanding this one.

CONCEPTUAL FRAMEWORK FOR TRAINING COURSE EVALUATION

At the end of a training course the training staff is faced with the question "Was this course successful?" As stated, the question is difficult to answer with any certainty. As with any complex event the answer to the larger question lies in the answers to a number of smaller questions.

It is noted in *Application of a Systems Approach to Training: A Case Study* that the eventual purpose of training is to provide learning experiences which influence trainee behavior according to predetermined objectives.[2] In an ideal sense, the demonstration of the desired behavior by the trainee constitutes direct evidence that the course objectives have been met and the training was successful. However, it is often necessary for the training specialist to use less direct evidence when attempting to determine the success of a training course. For example, it is easier to "demonstrate desired behavior" in courses designed to lead to concrete skills such as typing, card punching, and computation than in courses designed to lead to attitudinal changes, such as in supervisory and executive development courses.

This difference in ease of evaluation is well known to training specialists and results from the desire to measure the extent to which objectives are achieved and to interpret the results in terms of some standards. The specificity with which the objectives and standards can be established or defined will in large measure affect the extent to which measurement can occur. In general, the

See footnotes at end

more specific they are the easier it will be to evaluate and maintain control over the quality of the training product.

Quality control describes a concept developed in industry to assure that the end-product conforms to predetermined standards of quality, workmanship, size, and form. The basis of quality control, is a system of checks and inspections that will provide continuing feedback on the status of the manufactured product. Inspection occurs at critical points in the receipt, handling, and processing of materials as they flow through the plant. Engineering drawings, materials, specifications, and other similar criteria are used as a basis for determining quality in the inspection process.[3]

The three steps necessary to establish a quality control system are: Identification of control areas; determination of standards of quality; and systematic evaluation to ascertain whether or not standards are being met. With appropriate modifications these steps are applicable to the training process, as shown below:

Step 1. ***The Identification of Control Areas.***

The term "control areas", as used here, is identical to the establishment of training objectives.

Step 2. ***Determination of Standards of Quality.***

This is equivalent to the determination of what is to be considered acceptable achievement on the part of trainees: the standards of training performance.

Step 3. ***Systematic Evaluation.***

This involves the determination of whether standards are being met, which equates to (1) construction of instruments for assessing trainee progress, (2) periodic administration of these instruments (3) analysis and interpretation of data obtained and (4) feedback of results to training specialists and agency management for use in making needed course adjustments.

The ultimate purpose of quality control is to insure—through early identification and prompt correction of deficiencies as they arise during the processing or training period—an end product (manufactured article or trained person) which meets a predetermined standard of quality or proficiency. In the case of a training course, such continuous monitoring is necessary because even the most thorough planning and design represents only a best estimate of what will happen when the course is actually conducted.

This analogy, between industrial and training quality control, should not be carried too far. In the case of manufactured articles, the goal is to make them, insofar as possible, identical with a prototype. In the case of trainees, the goal is to bring each up to a predetermined minimum acceptable achievement standard; but beyond this, it would be expected that most trainees would exceed this minimum standard in varying degrees. The ideal of uniformity, therefore, would not apply in the training situation.

DEVELOPING TRAINING OBJECTIVES AND STANDARDS

Training objectives are goals to be reached as a result of the training course. Ideally, they should be stated in performance terms, i.e., some actions or behaviors which the trainee is expected to exhibit as a result of his training experience. It is best to have both interim and terminal training objectives, because this allows for periodic in-course, as well as end-of-course, evaluation.

The essential steps in formulating training objectives include:

—Determining what general improvements or changes are desired.

—Selecting out from the total improvements or changes desired those which would best be learned in the training course.

—Identifying the particular skills and knowledges which will lead to the desired performance.

—Developing training objectives by describing the action or behavior expected of the trainee as a result of the training.

When writing training objectives it is important to describe not only the action or behavior expected but also the conditions under which the behavior will be exhibited and the criterion, or standard, of acceptable training performance. For example, two training objectives illustrating these three elements might read:

See footnotes at end

Example 1.

The trainee will develop a program in COBOL which will print an alphabetic listing of those employee development specialists in the test deck who are at grade GS–11 and above. (*This describes the act to be performed.*) The trainee will work independently and, with the aid of his notes and other reference material, complete his assignment within 1 class day. (*These are the conditions.*) The trainee's program will be judged satisfactory if it runs free of error after no more than three compilations. (*This is the training performance standard.*)

Example 2.

Given appropriate reference materials and the format guide, and with minimal supervision (*These are the conditions.*), the instructor-trainee will demonstrate his ability to develop, in standard form, an instructor lesson plan suitable for a 30-minute classroom presentation. (*This is the act.*) The plan must be logical and properly detailed for platform use, and must be at least 90 percent correct as measured by the standard lesson plan checklist. The trainee will be given up to two hours to complete this task. (*This is the standard.*)

Since the subject matter of most courses is less precise than that represented by the two examples above, it is often difficult to state training objectives in such specific terms. Nevertheless, it is important for evaluation purposes to write training objectives as specifically as the subject matter permits. Here are two examples of training objectives written for less precise subjects:

Example 3.

Using the seven fundamentals of position classification as taught during the course (*These are the conditions.*), the trainee will demonstrate his skill at conducting a desk audit (*This is the act.*) which receives a rating of at least "Average" on the standard rating form completed by the instructor. (*This is the standard.*)

Example 4.

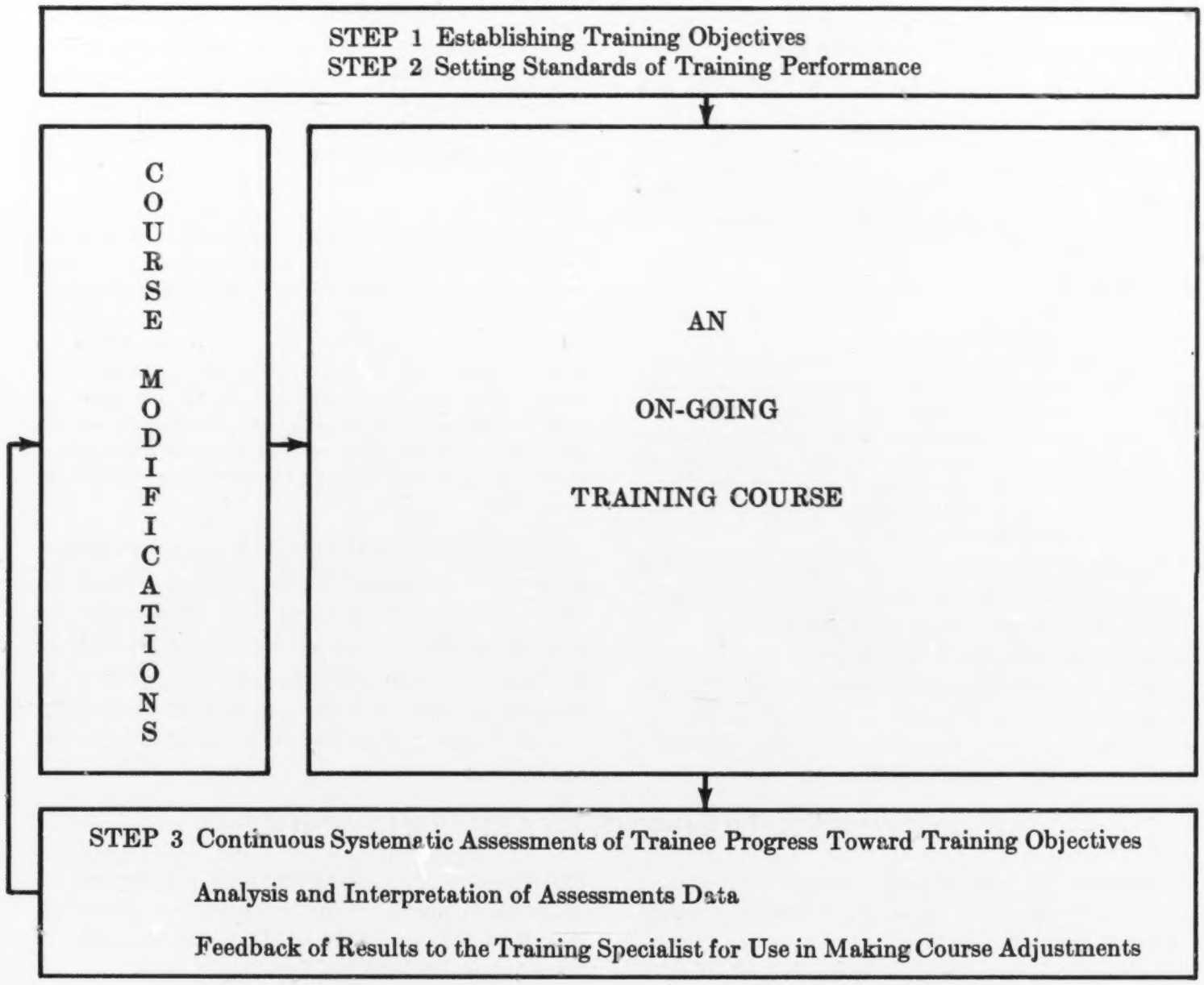
Given an employee relations problem in supervision, the trainee will demonstrate his ability to solve the problem using the six basic principles of problem solving taught in the course. (*This is the act.*) He will work independently and may use class notes and other course references as desired, and may have up to 50 minutes to complete his work. (*These are the conditions.*) To be considered acceptable, the trainee's solution must incorporate all six principles and result in a course of action which is judged feasible for implementation by the instructor. (*This is the standard.*)

Writing training objectives in such specific language will require much time, study, and perseverance. However, it is clear that when they are written in this form, subsequent tasks such as developing course content, selecting methods of presentation, and establishing criteria for quality control can be approached with considerably more confidence.

AN IN-COURSE TRAINING EVALUATION PROCESS

Once the training objectives have been written and the standards of training performance established the next step in the development of an evaluation design is to select the approaches which seem best fitted to the assessment of trainee progress. An in-course training evaluation process, incorporating all three steps of the quality control system in sequence, is displayed below.

FIGURE 1


STEP 1 Establishing Training Objectives
STEP 2 Setting Standards of Training Performance
COURSE MODIFICATIONS
AN
ON-GOING
TRAINING COURSE
STEP 3 Continuous Systematic Assessments of Trainee Progress Toward Training Objectives
Analysis and Interpretation of Assessments Data
Feedback of Results to the Training Specialist for Use in Making Course Adjustments

In the figure above, STEP 1 represents those actions or behaviors which the trainees are expected to exhibit. STEP 2 represents the minimum acceptable level of achievement for each training objective identified in STEP 1. Stated a little differently, STEP 1 represents the desired outcomes of the training course, and STEP 2 defines those points on the learning continuum at which the outcomes can be said to have been achieved.

STEP 3, then, represents that part of the process through which the training specialist determines, on a continuing basis, whether or not the trainees are meeting the standards established in STEP 2.

It should be noted that STEP 3 does not operate independently, or after the fact; rather it is part of a continuous feedback loop which contains four major components:

(1) An on-going training course designed to accomplish the objectives established in STEP 1;

(2) Continuous, systematic collection of data bearing on trainee achievement;

(3) Analysis and interpretation of data collected in terms of standards established in STEP 2; and

(4) Course modifications as appropriate.

Thus STEP 3 provides a continuing flow of in-course evaluative data to the training specialist about the status of each trainee. While the evaluative data received in STEP 3 is ordinarily used in making day-to-day adjustments in the on-going

program, there are circumstances where the adjustments required go beyond that and bring into question the appropriateness of the objectives or standards. When this occurs the feedback loop would also connect with STEPS 1 and 2.

Simply stated, in-course evaluation data can be classified into two types:

(1) Data that provides *direct*—or concrete—evidence about the extent to which the trainees can exhibit the desired actions or behaviors stated in a training objective; and

(2) Data that provides *indirect*—or inferential—evidence about training objective achievement.

Neither of these types of data taken alone can serve as a completely satisfactory basis for course evaluation. Both are needed to help counteract the limitations that are present in each. By collecting both types of data the training specialist will be in a much better position to ascertain whether or not course objectives are being met.

Direct Data About Trainee Achievement

This can be generated either informally or formally. Informal approaches such as role playing, case studies, class discussion, and trainee answers to instructor questions can provide the training specialist with an indication of the extent to which a trainee has achieved a given training objective. In some courses they are the principal source of direct data.

Nevertheless, when using informal methods it is often difficult to collect sufficiently reliable data about the progress of each trainee. To obtain such data it is necessary to use more formal methods of direct data collection, such as written tests or performance exercises. The principal advantage of direct evaluative data collected through formal sources is that it is the best source of concrete evidence about the progress each trainee has made toward achieving the training objectives.

The process of collecting direct data through formal sources has several potential limitations:

(1) The subject matter may not easily lend itself to formal measurement.

(2) Constructing direct measurement instruments is a difficult task requiring expertness that may not be readily available to the training specialist.

(3) Many trainees are embarrassed by "tests" and resent their use in training programs.

While these limitations are not inconsequential, the offsetting value of this source of achievement data often makes it worth the effort to overcome the limitations.

Indirect Data About Trainee Achievement

This may be as informal as a trainee's incidental remark after class or as formal as a survey of trainee opinion.

Informal data collection goes on continuously at the instructor level and is useful for suggesting minor adjustments in program presentation and administration. However, since such data is difficult to collect systematically and interpret with confidence, informal data can rarely provide conclusive evidence about trainee achievement.

As with direct data collection, the most reliable data is obtained through the use of formal methods such as trainee opinion surveys and classroom observations.

The advantage of *trainee opinion* data, when systematically collected and properly interpreted, derive from the fact that a trainee is usually present for the total course and is thus able to make personal judgments about several important aspects of it, such as:

(1) The extent to which he believes he has achieved the training objectives.

(2) Those factors which he believes were responsible for objectives being incompletely achieved.

(3) The applicability of what he has learned to what he believes he will need in the work situation.

(4) Whether or not he believes that the benefits received were worth the time and effort required.

In addition, such data provides a general summary of opinion about the overall quality of the course and the trainees' satisfaction with what they have learned, both of which can be compared with the opinions of other groups which have attended the same (or a similar) course.

However, data gathered from opinion surveys has two limitations which restrict the use to which that data can be put:

(1) Trainee opinions provide inferential rather than direct evidence about the extent to which training objectives have been met.

(2) It is difficult to assess the reliability and validity of the opinions expressed.

A complementary method for collecting *indirect* data about trainee achievement is classroom observation. Applying this method requires the use of observers who know the objectives and subject matter of the session being observed and are familiar with good training practices.

When systematically collected by qualified observers, the most important advantage of classroom observation data is that it provides third-party evidence which can be used to verify data derived from other sources.

The limitations of the classroom observation method are that:

(1) It represents only an individual's point of view. The observer, no matter how well qualified, cannot assess the instructional experience through the eyes of the learner.

(2) It may lack perspective. The observer who assesses an isolated part of a larger instructional block does not have firsthand knowledge of what occurred in prior or subsequent sessions.

DEVELOPING AND USING DIRECT MEASUREMENT INSTRUMENTS

Developing instruments for collecting data about trainee progress requires substantial investments of time and effort on the part of the training specialist.

To be effective, such measuring instruments must possess two fundamental qualities: validity and reliability.

Validity refers to the relevance of the instrument to the purpose for which it is being used. Does it in fact do what it is intended to do? Will results provided by the instrument serve as an accurate indicator of the extent to which established training goals have been achieved?

Reliability refers to the dependability of the data. How much reliance can be put in it? If data about the same (or comparable) classes were to be collected successively, would the results be substantially the same? If the instrument were to be used to gather data about two trainees who had achieved equally, would the results reflect this?

While the importance of developing valid and reliable measurement instruments is apparent, good instruments in themselves do not assure effective evaluation; slipshod use can detract substantially from their value in the training quality control process.

This section (and the one that follows) describe how to develop and use direct and indirect measurement instruments for collection of evaluative data.

The Development of Direct Measurement Instruments

This involves ten steps which, when followed by the training specialist, will facilitate the development of valid and reliable instruments. The same steps should be followed for the development of either written tests or performance exercises.

(1) **Decide what is to be measured.**

Accomplishing this step involves: (a) Selecting those training objectives against which trainee progress can directly be measured; and (b) Relating each objective to the instructional material which is intended to produce the desired action or behavior.

As noted earlier, this step is fundamental to the entire evaluation process.

(2) **Develop a measurement matrix.**

In simple terms, a measurement matrix is a worksheet which allows the training specialist to conveniently sketch out evaluation strategies which he believes will result in the most efficient and effective measurement of the extent to which training objectives have been met by the trainees. A sample measurement matrix is shown in figure 2. The seven major headings are discussed in the narrative text which follows. When completed, a simple measurement matrix would resemble the one shown in figure 3.

FIGURE 2

TRAINING MEASUREMENT MATRIX

Course Title:_____
Course Length:_____ Time to be Devoted to Measurement:_____

(1) Training objectives	(2) Hours of instruction	(3) Learning levels	(4) Measurement approach	(5) Assigned weights	(6) Estimated time Requirements	(7) Remarks

Column 1. *Training objectives* have previously been discussed in detail and, as noted in step 1 above, they are fundamental to the entire evaluation process.

Column 2. *Hours of instruction.* For convenience, total clockhours are normally rounded to the nearest whole hour.

Column 3. *Learning levels* can be conveniently categorized under three headings:

(*a*) Familiarization, where trainees learn sufficient facts and principles to be able to recognize their importance and know where to go for additional knowledge should the need arise.

(*b*) Working knowledge, where trainees possess a sufficient knowledge of related facts, principles, and techniques to enable them to perform routine practical applications under the direct supervision of a qualified individual.

(*c*) Qualified, where the trainees are capable of demonstrating sufficient knowledge to permit independent functioning in the area concerned.

Column 4. *Measurement approach.* Basically, there are two approaches by which trainee achivement can be directly measured. These are:

(*a*) Written tests, of which there are two main types: the objective (true/false, multiple choice, matching, etc.) and the subjective (essay); and

(*b*) Performance exercises, in which the trainee is required to perform the actual skills being taught under closely controlled conditions. Performance exercises are used frequently in the training of maintenance and service personnel. However, in recent years, such exercises have been used to train supervisors, counselors, and classifiers.

Naturally the approach selected will have to be congruent with the learning level identified in column 4. Thus:

- Measurement at the familiarization level would normally require trainees to recognize the correct facts and principles and suggests the use of objective (limited response) questions in the measuring instrument.
- Measurement at the qualified level would normally require extensive use of performance exercises.

From the above, it should not be inferred that there is an arbitrary relationship between the learning level and the type of approach used. Depending on the situation, satisfactory measurement at the working knowledge level, for instance, may be accomplished using objective test items. The important thing to remember is that the measurement approach employed should be compatible with the learning level.

Column 5. *Assigned weights.* Weights are assigned, in percentages, according to the importance of each training objective. If the number of hours of instruction given to accomplish each course objective were exactly equivalent to their importance, the weighting would be directly proportional. However, this is rarely the case. In most instances the amount of time required to bring about the desired objective will be either greater or less than the absolute weight (or importance) of that behavior in the total course.

Column 6. *Estimated time requirements* are educated guesses about the amount of time which would be required to administer the measurement instrument selected in column 5. A tally of the time requirements listed in column 6 for a given course can serve as a quick check on the feasibility of fitting the evaluation strategies selected to the course time schedule.

Column 7. *Remarks.* This column conveniently allows the training specialist to "footnote" his decisions with questions, qualifications, reminders, and suggestions.

FIGURE 3

TRAINING MEASUREMENT MATRIX

Course Title: _____

Course Length: _____ Time to be Devoted to Measurement: _____

(1) Training objectives	(2) Hours of instruction	(3) Learning level	(4) Measurement approach	(5) Assigned weights	(6) Estimated time requirements	(7) Remarks
				Percent	*Minutes*	
1 2 3	4	Familiarization	Objective written test	15	25	See if test includes items on new topics.
4 5 6 7 8	20	Qualified	Practical exercise	50	90	Use problem solving items.
9	8	Working	Objective written test	20	30	Use 30 true/false items.
10 11 12	3	Familiarization	Objective written test	10	20	Try changing 5 of the items to multiple-choice format.
13 14	2	Working	Subjective written test	5	15	Use essay format.
	37			100	180	

(3) **Develop the elements of the measurement instrument.**

The basic element of the written exercise is the test item. For the performance exercise the requirement would be to perform a required task. When developing measurement elements, several points should be kept in mind:

- The element employed should be compatible with the learning level as specified in the training objective statements.
- Elements should require the application of the information learned rather than being limited to mere recall of facts.
- Elements should be clearly stated so there is no chance of misinterpreting their intent.
- Only elements which bear directly on the attainment of training objectives should be used.
- Elements should range in difficulty from those which nearly all trainees may be expected to answer to those which challenge even the best trainees.

(4) **Assemble the measurement instrument.**

The following are generally accepted principles for arranging the elements of a measurement instrument:

- If different kinds of elements are used, group like kinds together.
- Insofar as possible, arrange the elements so that a few of the easier ones come first. This will give the trainee early success and not frustrate him.

(5) **Write directions for administering the instrument.**

Directions should be so complete that no additional information need formally be provided before, during, or after the exercise. In the case of a written exercise, the directions would appear on the cover sheet.

(6) **Dry run the instrument.**

This is normally done by having several members of the staff take the test or work through the exercise, and then obtaining their opinions on such factors as validity, practicability, organization, level of difficulty, and length.

(7) **Establish the critical achievement point.**

This entails the determination of what represents minimum acceptable achievement on the test or exercise. The establishment of the minimum satisfactory achievement level for a measuring instrument is a decision made after consideration of all relevant factors, such as length of exercise, level of difficulty, and performance requirements. Scales used for indicating achievement range from very coarse two-point scales, such as "satisfactory" and "unsatisfactory", to precise percentage scales (1 to 100). The scales employed would be dictated by the degrees of precision desired (or possible) in assessing achievement.

(8) **Administer the instrument.**

If the directions are complete and clearly presented, administration should go smoothly. If, on the other hand, trainees continually ask for clarification and additional information after the test or exercise is underway, some revision and/or expansion of directions is in order.

(9) **Analyze and evaluate results of the instruments.**

Although the preceding steps do much to assure the development of a good measuring instrument, they do not guarantee a finished product. Certain inadequacies may have slipped by unnoticed. Now, after it has been administered, results should be analyzed and inadequacies identified.

(10) **Revise the instrument based on the above analysis.**

This revision should begin immediately after weaknesses have been identified and verified.

Using Direct Measurement Instruments

This involves a relatively straightforward procedure consisting of three steps: Administer the instrument; process the results; and hold a critique. The following paragraphs describe how the steps should be applied to realize the full benefits from direct measurement instruments.

(1) **Administer the instrument.**

In order for a trainee to carry out a written or performance exercise, he must be fully informed about the parameters of the exercise. Directions should be clear and complete—the goal being to anticipate and answer those questions which a trainee might be expected to ask before the exercise begins. Directions normally include information on the following:

(*a*) Time considerations, including:

- Total time allocated to the exercise as well as allocations for subelements of the exercise where applicable; and

- Advice to the trainees on how to use their time effectively.

(*b*) Specific directions on how to carry out each different element in the exercise.

(*c*) Whether or not the trainee will be permitted to ask questions during the exercise or make use of reference materials.

(2) **Process the results.**

It is important to score and analyze the data collected quickly while trainee interest is still centered on the measured area of instruction.

(*a*) *Scoring.* Where measurement is accomplished by use of objective test items, accurate scoring presents no problem. Where other kinds of items are employed (problem solving and essay, for instance) the element of judgment is introduced and scoring can become quite subjective unless specific criteria are established for assigning credit. This caution applies equally to performance exercises.

(*b*) *Analysis of results.* This involves the graphic display of score distribution and the determination of mean, median, and modal scores. These data make possible a quick comparison of the achievement of this group with established norms (where an equivalent pretest has been used, the training specialist can also determine the progress which the trainees have made from their pretraining performance level). In addition, an item analysis, which shows group achievement on the individual items or parts of the exercise, should be done to pinpoint areas of trainee difficulty.

(3) **Hold a critique.**

In keeping with the learning principle of "immediate confirmation", a critique should be conducted immediately after the exercise has been scored and the results analyzed. A prolonged lapse of time between measurement and critique may cancel out the value of the latter as a teaching device.

When conducting the critique, the training specialist should concentrate on those parts of the test or exercise with which a significant portion of the class had difficulty. Seldom missed parts may be largely disregarded. To keep the critique from becoming a "face off" between the instructor (defending "school solutions" to elements selected for review) and individual trainees (who disagree with these solutions), a useful technique is to divide the class into committees and assign one or more of the exercise problems to each committee which then assumes the responsibility for reaching a consensus concerning the solution of their assigned problems and for presenting their solutions to the class. Such a trainee-centered critique can provide a more dynamic learning situation for trainees, and solutions thus presented may be more credible to those who missed the problems.

DEVELOPING AND USING INDIRECT MEASUREMENT INSTRUMENTS

This section will describe how to develop and use two kinds of indirect measurement instruments: those designed to gather trainee opinion and those for collecting the opinions of qualified classroom observers.

The two forms for indirect performance measurement described in this section are not meant to be prescriptive models. They are only sample forms which have been developed to capitalize on the values of opinion while minimizing the limitations of this source of information.

In general, the same steps described earlier for the development of direct measurement instruments apply here, with two exceptions: there is no need to develop a measurement matrix and no need to establish critical achievement points.

Developing a Trainee Opinion Survey

This involves the use of two types of questions: controlled or free response.

A controlled response question is one in which the responses are provided and the trainee checks the one which best reflects his opinion. This type of item has the advantage of forcing trainee responses toward specific issues. It also facilitates tabulation and analysis. The principal limitation to this approach is that the trainee must confine his response to the alternatives given and is not permitted to provide additional opinions which may be, in fact, more important than those asked by the survey designer.

A free response item is one that poses an open-ended question which the trainee answers in his

own words. The advantages and limitations of free response items are just the reverse of those for limited response.

- They are easy to develop, and allow for unlimited responses.
- The responses may be difficult to tabulate and analyze, and they may not relate to significant issues.

Typically, training specialists have found that a combination of the two types of items is preferable.

Figure 4 shows the front side of a sample questionnaire which solicits trainee opinion primarily in the area of objective attainment. The emphasis is, therefore, on instructional "effects" rather than causal factors. Opinions about these factors—instructor qualities, training aids, facilities, etc.—is asked for only when the trainee feels they prevented the intended instructional "effect" from being fully attained. Thus, the approach used in this form differs from the traditional approach, where trainees are requested to rate directly the elements of course delivery (or causal factors), and where, based on the resulting concensus, the trainer makes certain assumptions about the degree to which the desired instructional "effects" (or objectives) were attained.

Column 1 in the figure is keyed to a list of training objectives. This provides the trainee with a precise frame of reference within which to base his opinions about the extent to which each objective has been met.

Column 2 provides the trainee with an opportunity to register his opinions about the extent to which he personally believes that he has achieved each objective.

Column 3 allows the trainee to indicate those factors in the course which he feels impeded his achieving a given training objective.

Column 4 provides an opportunity for the trainee to look beyond the training situation to the work environment and make judgments about the degree to which he believes a given objective relates to his job. In effect, the trainee is asked to make judgments about the validity of the training objectives. (Where the training is designed to prepare him for a job which he has never performed, this column can be eliminated.)

Column 5 asks the trainee to indicate his opinion about the extent to which a given objective was worth the time and money expended on it during the course. (As with column 4 above, this column should be dropped where the trainees have no basis for judgment.)

The back of the form (not shown) can be used to solicit additional opinions of particular interest to the training specialist. These might include opinions about:

- General matters, such as the significant strengths and weaknesses of a course or the amount of emphasis placed on particular curriculum areas;
- Specific aspects of a course, such as the convenience of the training site, length of sessions, or use of a particular instructional aid;
- Personal changes in beliefs or attitude which do not lend themselves to performance measurement.

The Use of Trainee Opinion Surveys

This parallels the direct evaluation procedure with the exception of the critique, which is not applicable.

(1) **When should an opinion survey be given?**

Traditionally trainee comments have been sought at the end of a course. Survey questionnaires are distributed on the last day to be either completed in class or taken home and returned later. The two main advantages of conducting the survey at this time are the ease of administration and the perspective possessed by the trainee who, having experienced the entire course, is able to comment on how effectively its parts fit together in light of the overall objectives. A limitation to the survey at the end of the course is that the trainee may concentrate his comments on the last part of the course (which he remembers more vividly) and overlook the first part.

If the types of skills, knowledges, or attitudes being taught can be broken down into major course units, each governed by a performance objective, trainee opinion surveys can be administered at the end of each unit when the objectives and the associated instruction are fresh in the trainee's mind.

The decision as to which approach, a single end-of-course survey or several end-of-unit surveys, should be used depends largely on two factors. One factor is course length: If it is a short course, one survey at the end will likely suffice; while, if the course is long, it may be desirable to administer one or more interim end-of-unit surveys. The other factor relates to the number of times the

FIGURE 4

TRAINEE OPINION QUESTIONNAIRE

In the continuing effort to make this a more useful and meaningful training experience, we would appreciate your opinion about the extent to which you feel that you have attained the training objectives established for this block of instruction. Attached is a list of training objectives numbered 1 thru_____. Please respond to the questions below by placing a check mark in the appropriate block. In instances where your response is not entirely favorable, explain why in the space provided at the bottom and on the back of this sheet. Be completely frank in checking your reponses and making your comments. Attach extra sheets if additional writing space is required.

(1)	(2)			(3)											(4)			(5)		
Major training objectives (refer to objective statements)	As far as you are concerned, was this objective achieved?			If the objective was not at all, or only partially, achieved, what factors were responsible? (Please check and explain below)											Do you believe that this objective was related to the requirements of your job?			Do you believe the benefits derived from pursuing this objective were worth the time and effort?		
	Definitely	To some extent	Not at all	The instructor	Lesson content	Instructional methods	Instructional level	Time allocation	Test and practical excersises	Reference materials	Training aids	Classroom facilities	Other		Definitely	To some extent (explain below)	Not at all (explain below)	Definitely	To some extent (explain below)	Not at all (explain) below)
No. 1																				
No. 2																				
No. 3																				
No. 4																				
No. 5																				
No. 6																				
No. 7																				

Explanation of factors checked above_____

_____

_____

course has been conducted: If it has already been given several times, a single, end-of-course survey may well be all that is required. If, on the other hand, it is a new course, the training specialist may want to administer several end-of-unit surveys as a means of gaining more specific trainee comments and suggestions that can be referred to in making subsequent course revisions.

(2) Who should give it?

Because objectivity is important to the quality control effort, it is best to have the survey conducted by someone who has not been directly connected with the presentation of the course. This person, however, should be knowledgeable of course objectives as well as familiar with the tactics and strategies used to present the course material.

(3) How should it be given?

Regardless of the time chosen to conduct the survey, the three basic functions of the administrator are:

- To explain the purpose of the survey, in order to avoid confusion and thus increase the likelihood of receiving relevant comments.
- To provide an overview of the course and its objectives to help refresh the trainees' memories.
- To administer the survey.

(4) What should be done with the results?

Data, once received, should be promptly tabulated and analyzed by the trainer. In his analysis he should look primarily for trends rather than isolated opinion, and he should consider the narrative comments along with the ratings. Interpretation, including verification of findings by comparison with data from other sources, is the final step in this process.

Regardless of the care with which survey results are analyzed and interpreted, they can have only limited value unless they lead to program improvement. Therefore, once the commitment to adjust the course has been made, prompt follow-up action is essential for effective quality control.

Developing a Classroom Observation Approach

This involves procedures parallel to those used in developing the trainee opinion survey discussed above. Therefore, it must be designed to collect information about the extent to which training objectives are being met, and identify those factors in the instructional process which appear to impede or facilitate trainees' attainment of the objectives. Figure 5 shows a sample classroom observation guide, the main parts of which are:

1. *Training objectives.* These provide a frame of reference within which the observer will base his judgments about the extent of objective attainment.
2. *Assessment.* Here the observer indicates the extent to which the objectives appear to have been attained by the trainees.
3. *Checks on attainment.* Here the observer is required to identify and judge the effectiveness of the checks made by the instructor during the session to monitor the extent to which training objectives are being attained.
4. *Factors influencing objective attainment.* The observer is required to add specificity to the general assessment which he made in 2 above.

The Use of the Classroom Observation Approach

This differs from other methods in that the emphasis is on an individual observing the training situation and making informed judgments about the extent to which the objectives are being met. If the observation is to contribute significantly to the evaluation effort, the observer should have a sound background in the subject matter of the session he is monitoring, a precise knowledge of its training objectives, and an awareness of how this session ties into the entire instructional program.

The observer's presence during class should not be conspicuous. Therefore, the sample observation form shown (fig. 5) has been designed to allow the observer to check the appropriate blocks during the session and fill in the required explanation later.

After the session, the observer should discuss with the instructor those factors which were believed to have facilitated or impeded the attainment of training objectives. Where appropriate, the oral critique should be followed by a written report in which effective and ineffective aspects previously discussed are noted along with recommendations for improvement. Such reports assume additional importance after the trainee performance data has been obtained, since performance data only indicates *what* has occurred while the observation reports along with opinion surveys can be used to explain *why* the trainees did or did not meet the objectives.

FIGURE 5

CLASSROOM OBSERVATION GUIDE

Course _____ Session _____

Time _____ Date _____ Instructor _____

1. ***Training objectives:*** _____

2. ***Assessment:*** To what extent did training objectives appear to be attained? (Adverse factors should be explained in item 4 below.)
   a. ☐ Fully. (Minor deficiencies, if any.)
   b. ☐ Partially. (Some factors preventing full achievement.)
   c. ☐ Hardly at all.

3. ***Checks on attainment:*** What steps did the instructor take to assure himself that training objectives were being attained? _____

   How effective were they? _____

4. ***Factors influencing objective attainment:*** Factors that had an adverse effect on training objective attainment as indicated in 2 above. (Check any such factors and explain on the back of this sheet.)

Factor	*Check*	*Factor*	*Check*
A. Introduction to training session	____	I. Instructor's ability to direct discussion	____
B. Content	____	J. Appropriateness of material to group	____
C. Clarity of explanations	____	K. Time/material relationship	____
D. Instructor's apparent knowledge of subject	____	L. Individual trainee involvement	____
E. Subject matter organization	____	M. Instructional methods used	____
F. Selection and use of training aids	____	N. Reference materials	____
G. Instructor's manner (tact, etc.)	____	O. Classroom facilities	____
H. Instructor's ability to maintain trainees' interest	____	P. Tests and practical exercises	____
		Q. Other (specify)	____

_____
(Observer's Name)

FACTORS THAT INFLUENCE THE DECISION TO EVALUATE

Earlier in the paper a process for the evaluation of an agency's internally developed training courses was presented. Readers are cautioned that it represents an ideal process and not one which should be considered completely applicable to all courses or to all parts of a particular course. The evaluation process should be considered as a continuum which ranges from *little* evaluation (or informal feedback) to *comprehensive* evaluation (which would include all four data collection methods described in this paper—opinion surveys, class observations, written tests, and performance exercises). Naturally, the points along this continuum relate to the number and kind of evaluation methods employed. The real question then becomes: How much evaluation does a particular course require? Or, to state it differently: What point on the continuum best represents its evaluation needs?

There are three variables which agency officials should consider when trying to decide how much evaluation a particular course requires. One of them—course content—strongly determines what can be done. The other two—learning level and course costs—suggest what ought to be done.

(1) ***Course content.***

This is particularly influential on how much measurement can be applied. Typically, direct measurement cannot be accomplished with the same precision in human relations training, for example, as it can in courses that are more technical in nature.

(2) ***The Learning Level of the Course.***

Is the purpose of the course essentialy to familiarize, provide a working knowledge, or qualify? A course designed to qualify trainees in a particular skill requires a more complete evaluation than an orientation program designed to familiarize entering employees with an agency's operating procedures.

(3) ***Course costs.***

In general, the more costly the course, the more comprehensive its evaluation ought to be. Costs may be considered from two points of view:

(*a*) Per unit (trainee) cost. Factors influencing this cost are: Number of trainees; course length; personnel, equipment, and facilities requirements for conducting the course; and trainee grade level.

(*b*) Course frequency. Courses that are repeated frequently may have a relatively low per unit training cost, but may represent considerable overall cost because of the large number of personnel trained. Such courses may deserve comprehensive evaluation. On the other hand, extensive evaluation of "one-time" courses is seldom justified.

A consideration of these variables helps to give agency officials a picture of what kind of evaluation plan a particular course may need. Whether or not such a plan can be implemented is influenced by two factors.

One of these factors is the degree of control exercised by the agency over course delivery. Where this control covers all elements—media, facilities and, in particular, the instructional staff—the level of evaluation may be set based on consideration of the above variables. However, where it is desirable to use frequent guest faculty to present the program, control over the delivery stage will be more limited. As a consequence, it will be more difficult to apply achievement measures.

The second influencing factor will be the resources available to the agency. A full scale evaluation plan represents considerable time, money, and professional competency on the part of the training personnel involved. Since increased reliability comes at a high price, it must be decided what point on the continuum constitutes an acceptable trade off between cost and reliability. This decision making is further complicated because it does not normally involve only a single course, but rather every internally developed course, making up the agency's training program. Agency officials must therefore set priorities and assign resources based on their needs for information—information which will tell them whether or not a given training experience has accomplished what was intended.

Footnotes

[1] U.S. Civil Service Commission, Bureau of Training, *Instructional Systems and Technology: An Introduction to the Field and Its Use in Federal Training*, Training Systems and Technology Series: No. I (Washington, D.C.: Government Printing Office, 1969), pp. 2–3.

[2] U.S. Civil Service Commission, Bureau of Training, *Application of a Systems Approach to Training: A Case Study*, Training Systems and Technology Series: No. II (Washington, D.C.: Government Printing Office, 1969), p. 1.

[3] U.S. Army Transportation School, *Educational Quality Control* (Fort Eustis, Va.: U.S. Army Transportation School, 1967), p. 1.

Selected Bibliography

The following selected list of references was prepared for the reader who desires to study in more detail the concepts and procedures contained in this paper. For convenience, the list is divided into specific areas addressed in the paper.

Training Quality Control

Continental Army Command. *Systems Engineering of Training*. Regulation No. 350–100–1. Fort Monroe, Va.: Continental Army Command, 1968.

Smith, Robert G., Jr. *Controlling the Quality of Training*. Technical Report 65–6. Washington, D.C.: Human Resources Research Organization, 1965.

U.S. Army Engineer School. *Quality Control Media*. Fort Belvoir, Va.: U.S. Army Engineer School, 1967.

U.S. Army Transportation School. *Educational Quality Control*. Fort Eustis, Va.: U.S. Army Transportation School, 1967.

Training Objectives

Bloom, Benjamin S., ed. *Taxonomy of Educational Objectives, Handbook I: Cognitive Domain*. New York: David McKay Co., Inc., 1956. Reprinted 1968.

Continental Army Command. *Student Performance Objectives*. Training Pamphlet No. 350–14. Fort Monroe, Va.: Continental Army Command, 1966.

Direct Measuring Techniques

Adkins, Dorothy C. "Measurement in Relation to the Educational Process." *Educational and Psychological Measurement, 18:2* (Summer, 1958), pp. 221–240.

Angell, David; James W. Shearer; and David C. Berliner. *Study of Training Performance Evaluation Techniques*. Technical Report NAV TRADEVCEN 1449–1. Palo Alto: American Institute for Research, 1964.

Bloom, Benjamin S. *Toward a Theory of Testing which includes Measurement—Evaluation—Assessment*. CSEIP Occasional Report No. 9. Los Angeles: University of California, 1968.

Collins, A. K. *Assessing the Effectiveness of Industrial Training*. Canberra, Australia: Commonwealth Public Service Board, Training Section, 1961.

Glaser, Robert. "Institutional Technology and the Measurement of Learning Outcomes: Some Questions." *American Psychologist, 18:8* (August, 1963), pp. 519–521.

Lindquist, E. F., ed. *Educational Measurement*. Washington, D.C.: American Council on Education, 1963.

Micheels, W. J., and M. R. Karnes. *Measuring Educational Achievement*. New York: McGraw-Hill Book Co., 1950.

Remmers, H. H.; N. L. Gage; and J. F. Rummell. *A Practical Introduction to Measurement and Evaluation*, 2d ed. New York: Harper and Row, 1965.

U.S. Army Engineer School. *Critique of Written Examination*. Lesson Reference File IG 011–5. Fort Belvoir, Va.: U.S. Army Engineer School, 1964.

U.S. Army Engineer School. *Guidelines for the Preparation and Administration of Graded Examinations*. Regulation No. 350–2. Fort Belvoir, Va.: U.S. Army Engineer School, 1965.

U.S. Army Southeast Signal School. *Training Evaluation Plans*. Fort Gordon, Ga.: U.S. Army Southeast Signal School, 1969.

Indirect Measurement Techniques

Adams, J. Stacy. *Interviewing Procedures.* Chapel Hill: The University of North Carolina Press, 1958.

Catalanello, Ralph F., and Donald L. Kirkpatrick. "Evaluating Training Programs—The State of the Art." *Training and Development Journal, 22:5* (May, 1968), pp. 2–9.

U.S. Army Engineer School. *Training Quality Control—Student Opinion Surveys.* Regulation No. 350–13. Fort Belvoir, Va.: U.S. Army Engineer School, 1970.

U.S. Army Engineer School. *Supervision of Instruction.* Lesson Reference File IG. 008–1. Fort Belvoir, Va.: U.S. Army Engineer School, 1967.

General

Alkin, Marvin C. *Towards an Evaluation Model: A Systems Approach.* CSEIP Working Paper No. 4. Los Angeles: University of California, 1968.

Hesseling, P. *Strategy of Evaluation Research.* Assen, The Netherlands: Van Gorcum, Ltd., 1966.

Kirkpatrick, Donald K. "How to Start on Objective Evaluation of Your Training Program." *American Society of Training Directors, 10:3* (May–June, 1956), p. 18.

U.S. Civil Service Commission. *Instructional Systems and Technology: An Introduction to the Field and Its Use in Federal Training.* Training Systems and Technology Series: No. I. Washington, D.C.: Government Printing Office, 1969.

U.S. Civil Service Commission. *Application of a Systems Approach to Training: A Case Study.* Training Systems and Technology Series: No. II. Washington, D.C.: Government Printing Office, 1969.

U.S. GOVERNMENT PRINTING OFFICE:1971 O—413-921